心理学报 2023, Vol. 55, No. 8, 1372−1382
Acta Psychologica Sinica

https://doi.org/10.3724/SP.J.1041.2023.01372

二阶潜增长模型标度方法及其可比的一阶潜增长模型*

温忠麟[1] 王一帆[1] 杜铭诗[1] 俞雅慧[1] 张愉蕙[1] 金童林[2]

([1]华南师范大学心理应用研究中心/心理学院, 广州 510631) ([2]内蒙古师范大学心理学院, 呼和浩特 010022)

摘 要 潜增长模型(LGM)是分析纵向数据的一种强有力工具, 在心理学和其他社会科学研究领域受到重视。多指标测量的变量, 既可以用合成分数建立单变量 LGM(一阶 LGM), 也可以用指标建立潜变量 LGM(二阶 LGM)。简述了二阶 LGM 标度方法(包括尺度指标法和效应编码), 提出了有可操作性的潜变量标准化标度方法和合成分数的一阶 LGM 标准化模型。系统总结了二阶 LGM 标度方法及其可比的一阶 LGM 建模, 并用多指标变量的实测数据进行示例。推荐使用效应编码法对二阶 LGM 进行标度和标准化。

关键词 潜增长模型, 合成分数, 标度方法, 尺度指标, 效应编码, 潜变量标准化

分类号 B841

为了模型识别(identification), 验证性因子分析(CFA)的一项基础工作是因子标度(scaling), 即为因子设定测量单位和测量原点(侯杰泰 等, 2004)。先说设定测量单位, 通常的标度方法(scaling method, 也称为识别方法)是使用固定负荷法或者固定方差法指定测量单位。固定负荷法是对每个因子, 都将其中一个指标的负荷(loading)固定为一个常数, 相应的指标称为尺度指标(scaled-indicator 或者 marker-variable)。最方便的是固定负荷为 1, 此时因子的测量单位与尺度指标的测量单位相同; 固定方差法是将每个因子的方差固定为一个常数, 最方便的是固定方差为 1。再说设定测量原点, 对于无均值结构的模型, 其实是将因子均值和指标的截距(intercept)都固定为 0; 对于有均值结构的模型(Mplus 软件默认), 有两种标度方法: 一种是将因子的均值固定为 0, 另一种是将尺度指标的截距固定为 0。

潜增长模型(latent growth model, LGM)是分析纵向数据的一种强有力工具, 在心理学和其他社会科学研究领域受到重视(例如, Jeon & Kim, 2021; Yang et al., 2021; 也见本文讨论部分)。LGM 是一种特殊的因子模型, 它的标度方法有什么特殊性? 用潜变量(latent variable)建立的 LGM 应当如何标度, 才能与用显变量建立的 LGM 有可比性? 这些是本文要研究的主要问题。本文先介绍单变量 LGM (一阶 LGM)和潜变量 LGM (二阶 LGM); 然后简述二阶 LGM 标度方法, 提出有可操作性的潜变量标准化标度方法, 并在此基础上系统总结二阶 LGM 标度方法及其可比的一阶 LGM 建模; 接着使用一个多指标变量的实测数据, 对所论的标度方法建模并比较一阶和二阶 LGM 的参数估计结果; 最后对二阶 LGM 标度和模型标准化方法做出推荐。

1 潜增长模型

1.1 单变量潜增长模型

虽然有一些文献从潜在状态−特质理论(latent state-trait theory)出发导出 LGM 方程(例如, Mayer et al., 2012; Geiser et al., 2013), 但更直接、更容易理解的还是从时间 t 为自变量的普通回归方程导出(例见: 李丽霞 等, 2012)。

本节考虑单变量的潜增长模型(univariate LGM),

收稿日期: 2023-02-06

* 国家自然科学基金项目(32171091)、教育部人文社会科学重点研究基地重大项目(22JJD190006)资助。

并列第一作者: 王一帆

通信作者: 温忠麟, E-mail: wenzl@scnu.edu.cn

是一阶因子模型的特例。相对于后面要讨论的二阶 LGM，也称为一阶 LGM。设变量 X 做了 T 次追踪测量，记为 X_1，X_2，⋯，X_T；再设时间间隔相等，时间点记为 $t = 1, 2, \cdots, T$。为了设定第 1 个测量值 X_1 作为初始值，用测量时间点 t 减去 1 作为时间变量。对每个被试 i，都有一个回归方程去描述变量 X 随时间的变化

$$X_{ti} = \alpha_i + \beta_i(t-1) + e_{ti}, t = 1,2,\cdots,T \tag{1}$$

其中截距 α_i 表示 X 的初始水平(initial level)，斜率(slope) β_i 表示 X 的变化率(change rate)，e_{ti} 是时间点 t 的残差，通常假设 e_{ti} 是均值为零的随机变量。

每个被试 i 都有相应的 α_i 和 β_i，但不能直接观测得到，属于潜变量(即因子)。熟悉因子分析的读者，不难将方程(1)理解为有两个因子(α 和 β，分别称为截距因子和斜率因子)、T 个指标($X_1, X_2, \ldots, X_T$)的测量方程。为了简便起见，下面以 $T = 4$ 为例，测量方程如下(按通常因子模型的写法忽略被试下标 i)：

$$X_1 = 1\cdot\alpha + 0\cdot\beta + e_1 \tag{2}$$

$$X_2 = 1\cdot\alpha + 1\cdot\beta + e_2 \tag{3}$$

$$X_3 = 1\cdot\alpha + 2\cdot\beta + e_3 \tag{4}$$

$$X_4 = 1\cdot\alpha + 3\cdot\beta + e_4 \tag{5}$$

其中 e_t 的均值为 0，彼此不相关，并且与 α 和 β 都不相关。

图 1 是一个 4 次测量的一阶 LGM 路径图。因子 α 和 β 的均值分别记为 μ_α 和 μ_β，α 和 β 的方差分别记为 σ_α^2 和 σ_β^2，α 和 β 的协方差记为 $\sigma_{\alpha\beta}$。这 5 个 LGM 参数后面会多次提到，下面根据方程(1)~(4)推导含有这 5 个参数的方程。

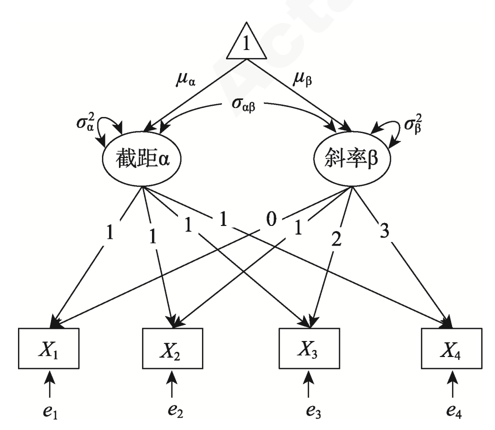


图 1　一阶 LGM 的例子

注：三角形出发的路径系数是与均值结构有关的参数。

方程(2)两边求均值

$$\mu_\alpha = E(\alpha) = E(X_1) \tag{6}$$

μ_α 是平均初始水平。由方程(3)减去(2)后两边求均值

$$\mu_\beta = E(\beta) = E(X_2) - E(X_1) \tag{7}$$

μ_β 是平均变化率。对于线性增长模型，μ_β 也就是 X 在前后两个时间点的平均差异。

由方程(2)~(4)，计算 X_1，X_2，X_3 之间的协方差可得

$$\text{cov}(X_1, X_2) = \sigma_\alpha^2 + \sigma_{\alpha\beta} \tag{8}$$

$$\text{cov}(X_1, X_3) = \sigma_\alpha^2 + 2\sigma_{\alpha\beta} \tag{9}$$

$$\text{cov}(X_2, X_3) = \sigma_\alpha^2 + 3\sigma_{\alpha\beta} + 2\sigma_\beta^2 \tag{10}$$

求解方程组(8)~(10)得到

$$\sigma_\alpha^2 = 2\text{cov}(X_1, X_2) - \text{cov}(X_1, X_3) \tag{11}$$

$$\sigma_{\alpha\beta} = \text{cov}(X_1, X_3) - \text{cov}(X_1, X_2) \tag{12}$$

$$\sigma_\beta^2 = \frac{1}{2}\text{cov}(X_2, X_3) + \frac{1}{2}\text{cov}(X_1, X_2) - \text{cov}(X_1, X_3) \tag{13}$$

σ_α^2 和 σ_β^2 分别反映初始水平和变化率的变异大小。$\sigma_{\alpha\beta}$ 反映变化率与初始水平的关系，如果正相关，则初始水平越高的被试，变化率往往越大；如果负相关，则初始水平越高的被试，变化率往往越小。此外，需要估计的参数是误差方差，而负荷已经全部固定。

熟悉多水平模型的读者或许已经看出来，μ_α 和 μ_β 是截距和斜率的固定效应，而 σ_α^2 和 σ_β^2 是截距和斜率的随机效应。有关多水平模型和潜增长模型分析纵向数据的比较，可参见 Chou 等人(1998)和李丽霞等人(2014)的文章。

LGM 一定要有均值结构，因为需要考虑指标均值和因子均值才能谈论增长。LGM 的负荷已经固定，因而一阶 LGM 的标度只涉及到设定因子的原点。因子均值在 LGM 中是增长曲线参数的固定效应，通常自由估计，因而因子标度的时候需要固定的是指标的测量方程的截距。首先，为了使得因子 α 表示初始水平，X_1 的截距 τ_1 必须为 0，否则由 $X_1 = \tau_1 + 1\cdot\alpha + 0\cdot\beta + e_1$ 可知，α 不等于初始水平。然后，X_2, X_3, X_4 的测量方程的截距都应当与上一个时间点的相同，否则前后测量的变化值会被测量方程截距的差异分走一部分，因子 β 不会刚好等于变化率，这从方程(3)~(5)加上不是 0 的截距项容易看出来。所以，LGM 中全部指标的测量方程的截距通常都设定为 0。

虽然本节讨论的是测量间隔相等的线性增长模型，但不难推广到一般的情形。当 X 随时间变化是非线性增长时，可以在回归方程中增加(t − 1)的二次甚至更高次的项，变成多项式回归，但多项式的次数必须少于追踪测量次数。例如，模型包含二次增长，则相应的 LGM 在截距、斜率(一次)因子之

外，需要增加一个二次因子，相应地，负荷矩阵增加一列，由$(t-1)$的平方产生，依次为 0, 1, 4, 9。包含更高次多项式的 LGM 依次类推。当时间间隔不相等时，将第一个间隔设定为 1，后面的间隔根据其与第一个间隔的长短比例进行调整，t 可能不是整数。例如，4 个测量时间点的间隔依次为 2 个月、3 个月、2 个月，则 t =1, 2, 3.5, 4.5，即 t 增加 1 表示间隔 2 个月。

1.2 二阶潜增长模型

在心理、行为、管理、教育等研究领域，变量往往不是显变量，而是需要一组题目去间接测量的潜变量(记为 η)。这时既可以使用题目的合成分数(composite score)去建立单变量 LGM，也可以使用潜变量(以题目为指标)去建立二阶 LGM (second-order latent growth model)。

直接使用(题目)指标建立潜变量测量模型，然后使用每个时间点的潜变量(也就是通常 CFA 中的因子)建立 LGM，就得到 McArdle (1988)所说的因子曲线模型(the curve-of-factors model)。这种二阶 LGM，其实是一种特殊的二阶因子模型，其中一阶因子是一组题目测量的潜变量，二阶因子才是截距因子和斜率因子。

不失一般性，假设潜变量 η 有 3 个指标，指标可以是题目得分或者题目小组打包后的合成分数(吴艳，温忠麟，2011)。时间点 t 的 3 个指标得分记为 x_{t1},x_{t2},x_{t3}，t = 1, 2, 3, 4。二阶 LGM 中的一阶因子测量方程为

$$x_{t1}=\tau_1+\lambda_1\eta_t+\varepsilon_{t1} \quad (14)$$

$$x_{t2}=\tau_2+\lambda_2\eta_t+\varepsilon_{t2} \quad (15)$$

$$x_{t3}=\tau_3+\lambda_3\eta_t+\varepsilon_{t3} \quad (16)$$

其中 t = 1, 2, 3, 4 (故测量方程共有 12 个)，τ_i 是第 i 个指标的截距，λ_i 是第 i 个指标的因子负荷，ε_{ti} 是时间点 t 的第 i 个指标的误差项(均值为 0，彼此不相关，并且与诸 η 不相关)。截距和负荷不随时间变化，是因为假设了纵向测量不变性(后面有进一步说明)。

二阶 LGM 中的二阶因子测量方程为

$$\eta_1=1\cdot\alpha+0\cdot\beta+\zeta_1 \quad (17)$$

$$\eta_2=1\cdot\alpha+1\cdot\beta+\zeta_2 \quad (18)$$

$$\eta_3=1\cdot\alpha+2\cdot\beta+\zeta_3 \quad (19)$$

$$\eta_4=1\cdot\alpha+3\cdot\beta+\zeta_4 \quad (20)$$

其中 α 和 β 的含义与单变量 LGM 的相同，ζ_i 的均值为 0，彼此不相关，并且与 α 和 β 都不相关。二阶因子 α 和 β 的测量方程不需要截距项，道理和一阶 LGM 中 α 和 β 的测量方程不需要截距项一样。

二阶 LGM 路径图见图 2。

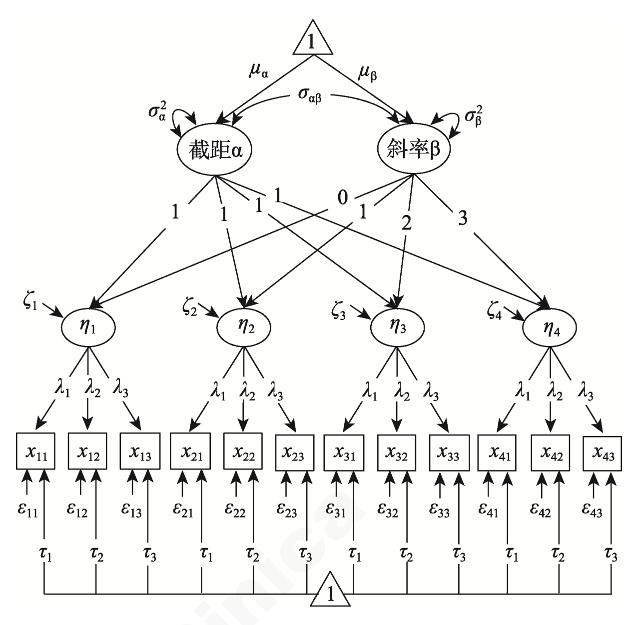


图 2 二阶 LGM 的例子

注：三角形出发的路径系数是与均值结构有关的参数。假设纵向测量强不变性成立，即指标负荷和截距不随时间变化。

由(17)和(18)可知，与方程(6)和(7)相应的是

$$\mu_\alpha=E(\alpha)=E(\eta_1) \quad (6')$$

$$\mu_\beta=E(\beta)=E(\eta_2)-E(\eta_1) \quad (7')$$

同理可得到类似于(11)~(13)那样的方程，其中的 X 替换为 η。

二阶 LGM 的一个重要前提假设是纵向测量不变性，并且要求强不变性(温忠麟，刘红云，2020)，即各个时间点的因子负荷和指标截距都不变(这里 3 个指标的负荷为 $\lambda_1,\lambda_2,\lambda_3$，截距为 τ_1,τ_2,τ_3，不随时间点变化)。如果不变性不满足，可能会给主要参数的估计带来偏差，偏差大小与违反不变性的参数(负荷违反、截距违反或者两者都违反不变性)、数量(违反不变性的负荷和截距的比例有多少)和程度(违反不变性的负荷或截距在各时间点的最大值与最小值之差)有关(Kim & Willson, 2014)，而且与潜变量的标度方法也有关(Jeon & Kim, 2021)。

2 二阶潜增长模型的标度方法

二阶 LGM 的一阶因子(即潜变量)和二阶因子(即截距因子和斜率因子)都需要标度。二阶因子的标度方法和单变量 LGM 的一样，这里重点讨论二阶 LGM 的一阶因子的标度方法。Yang 等人(2021)介绍了三种标度方法，包括尺度指标法、效应编码法和潜变量标准化法。无论哪种方法，都有一个等式限制负荷(设定潜变量测量单位)、一个等式限制

截距(设定潜变量测量原点)。下面我们重新演绎，简化了很多。尤其是对原来缺乏可操作性的潜变量标准化法，提出了简单而有可操作性的新方法。

2.1　尺度指标法

尺度指标法(scaled-indicator method)是最常用的方法：选定一个指标，将其截距固定为 0，负荷固定为 1。因为总可以将选定的指标排在第 1 个位置，所以下文就将第一个指标 x_{t1} 作为尺度指标。

在方程(14)中，设定 $\tau_1 = 0, \lambda_1 = 1$，有 $x_{t1} = \eta_t + \varepsilon_{t1}$，从而

$$E(\eta_t) = E(x_{t1}), t = 1,2,3,4 \tag{21}$$

再由方程(6')、(7')结合(21)可得

$$\mu_\alpha = E(x_{11}), \mu_\beta = E(x_{21}) - E(x_{11}) \tag{22}$$

由(21)和(22)可知，初始水平均值 μ_α 就是尺度指标的平均初始水平，变化率均值 μ_β 就是尺度指标在第 2 次测量与第 1 次测量的平均差异。对于线性增长模型，μ_β 也就是尺度指标在前后两个时间点的平均差异。除了 μ_α 和 μ_β 外，也可以将另外 3 个重要参数 $\sigma_\alpha^2, \sigma_\beta^2$ 和 $\sigma_{\alpha\beta}$ 用 x_{11}, x_{21}, x_{31} 之间的协方差表示出来(Yang et al., 2021)，结果与(11)~(13)的一样，只要将其中的 X_1, X_2, X_3 换成 x_{11}, x_{21}, x_{31} 便可。显然，一阶因子选用不同的尺度指标，得到的 LGM 参数估计结果可能是很不相同的。

这里的二阶线性 LGM 需要估计的参数包括：2 个负荷 λ_2, λ_3，2 个截距 τ_2, τ_3，12 个指标(4 个时间点合计)的误差方差，4 个时间点潜变量 $\eta_1, \eta_2, \eta_3, \eta_4$ (被截距因子和斜率因子解释后)的残差方差，5 个 LGM 参数(μ_α 和 $\mu_\beta, \sigma_\alpha^2$ 和 $\sigma_\beta^2, \sigma_{\alpha\beta}$)，合计 25 个。一般地，假设有 p 个指标，T 个时间点，则需要估计的参数包括：$p-1$ 个负荷，$p-1$ 个截距，pT 个指标的误差方差，T 个时间点的潜变量的残差方差，5 个 LGM 参数，一共有 $k = 2(p-1) + (p+1)T + 5$ 个。因为有均值结构的 CFA 模型的数据点包括 $pT(pT+1)/2$ 个样本协方差和 pT 个样本均值，减去自由估计的参数个数 k，便可得到二阶线性 LGM 的自由度的计算公式为

$$df_{2\text{LGM}} = pT(pT+1)/2 + pT - k = pT(pT+1)/2 - (2p+T+3) \tag{23}$$

后面介绍的两种潜变量标度方法得到的自由度与尺度指标法得到的一样。如果所写的程序运行结果得到的自由度与理论计算的不同，说明程序没有写对，很多时候都是潜变量标度有误。但要注意，如果不是线性 LGM，需要估计的参数一般会更多，自由度计算公式需要相应改变。

2.2　效应编码法

效应编码法(effect-coding method)是将各时间点的一阶因子测量方程的全部截距之和限制为 0，同时将全部负荷之和限制为一个常数(Little et al., 2006; Yang et al., 2021)，下面讨论文献上最常见的两种做法。

由方程(14)~(16)可知，对任一时间点 t，3 个指标之和为

$$x_{t1} + x_{t2} + x_{t3} = (\tau_1 + \tau_2 + \tau_3) + (\lambda_1 + \lambda_2 + \lambda_3)\eta_t + (\varepsilon_{t1} + \varepsilon_{t2} + \varepsilon_{t3}) \tag{24}$$

记 $X_t = x_{t1} + x_{t2} + x_{t3}$，它是由指标总和做成的合成分数。如果限制

$$\tau_1 + \tau_2 + \tau_3 = 0$$
$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{25}$$

则(24)变成 $X_t = \eta_t + \varepsilon_t$，其中 $\varepsilon_t = \varepsilon_{t1} + \varepsilon_{t2} + \varepsilon_{t3}$ 是合成分数的误差项。从而

$$E(\eta_t) = E(X_t), t = 1,2,3,4 \tag{26}$$

再由方程(6')、(7')结合(26)可得

$$\mu_\alpha = E(X_1) \tag{27}$$

$$\mu_\beta = E(X_2) - E(X_1) \tag{28}$$

方程(27)、(28)分别与方程(6)、(7)形式一模一样。同理还可推出(11)~(13)一样的方程组。这说明，二阶 LGM 中限制截距之和为 0、负荷之和为 1，得到的 5 个参数方程与使用合成分数(总和)建立单变量 LGM 得到的一样(但还有其他不一样的方程，因为 LGM 是超识别的，方程个数多于参数个数)，两者有接近的参数估计。

在方程(24)两边同时除以指标个数(这里是 3)，并记 $X_t = (x_{t1} + x_{t2} + x_{t3})/3$ (平均分做成的合成分数)，$\varepsilon_t = (\varepsilon_{t1} + \varepsilon_{t2} + \varepsilon_{t3})/3$ (合成分数的误差项)，则变为限制

$$\tau_1 + \tau_2 + \tau_3 = 0$$
$$\lambda_1 + \lambda_2 + \lambda_3 = 3 \tag{29}$$

得到与(26)一模一样的方程。这说明，二阶 LGM 限制截距之和为 0、负荷之和为指标个数，得到的参数方程中有 5 个与使用合成分数(平均分)建立单变量 LGM 得到的一样(但还有其他不一样的方程)，两者有接近的参数估计。

正如使用总分与使用平均分做单变量 LGM 得到的参数估计很不相同，限制负荷之和等于 1 与等于指标数，得到的二阶 LGM 参数估计也很不相同，两者的均值参数有倍数(指标个数)关系，方差也有倍数(指标个数的平方)关系。

2.3 潜变量标准化法

前面介绍的两种标度方法可能会得到很不同的二阶 LGM 结果，不仅不同标度方法得到的结果没有可比性，同一种标度方法下选用不同限制等式得到的结果也没有可比性。Yang 等人(2021)借鉴前人的做法(Ferrer et al., 2008; Grimm et al., 2017)，将某个时间点的潜变量进行标准化，变成均值为 0、方差为 1 的标准化变量。这个时间点称为参照点，如果没有特别说明，默认选择第一个时间点，相当于将初始水平设定为 0，初始方差设定为 1，并作为比较的基准，而后面时间点潜变量在纵向测量强不变性限制下相应变化均值(一般不是 0)和标准差(一般不是 1)。

Yang 等人(2021)提出的基于尺度指标法的潜变量标准化法(latent-standardization method)如下：

1)将尺度指标的截距固定为尺度指标的均值，即 $\tau_1 = E(x_{11})$；

2)将尺度指标的负荷固定为尺度指标的方差乘以其信度后开方，即

$$\lambda_1 = \sqrt{\sigma^2_{x_{11}} \rho_{x_{11}}} \tag{30}$$

下面解释一下其做法，以理解这种做法的本质。由方程(14)有

$$x_{11} = \tau_1 + \lambda_1 \eta_1 + \varepsilon_{11} \tag{31}$$

真分数的方差等于 $\mathrm{var}(\lambda_1 \eta_1) = \lambda_1^2 \mathrm{var}(\eta_1)$。令 $\mathrm{var}(\eta_1) = 1$，则真分数的方差等于 λ_1^2。注意到 x_{11} 的信度等于其真分数方差除以观测分数方差，即 $\rho_{x_{11}} = \lambda_1^2 / \sigma^2_{x_{11}}$，解出 λ_1 便得到(30)。

基于效应编码法也可以推出类似(30)的限制等式，只是尺度指标换成合成分数，信度换成合成分数的信度(Yang et al., 2021)。他们的示例中是假设已经知道了指标的信度，但实际应用中的信度是不知道的，因而他们提出的方法缺乏可操作性。如何知道指标或合成分数的信度呢？最可靠的是通过 CFA 得到，其他方法得到的信度(如重测信度、α 系数等)精确度较差(温忠麟，叶宝娟，2011；温忠麟等，2022)。信度不是 CFA 中的参数，而是通过 CFA 的参数计算得到。如果先做了 CFA，无需通过(30)去绕弯子计算负荷 λ_1，可以通过固定 η_1 的方差为 1、均值为 0，设置负荷和截距有跨时间的不变性(还可以顺便检验纵向测量强不变性是否成立)，得到的负荷就可以直接使用了。

这样，我们可以提出既简单又有可操作性的潜变量标准化法(两阶段建模)。基于尺度指标法的潜变量标准化法如下：

第一阶段：固定 η_1 的方差为 1、均值为 0，设置负荷和截距有跨时间的不变性，做 CFA；

第二阶段：建立二阶 LGM，将尺度指标的截距固定为第一时间点的尺度指标的样本均值，即 $\tau_1 = E(x_{11})$；将尺度指标的负荷 λ_1 固定为上面 CFA 得到的负荷。

基于效应编码法的潜变量标准化法如下：

第一阶段与基于尺度指标法的潜变量标准化法一样；

第二阶段：建立二阶 LGM，将各指标的截距之和限制为第一时间点各指标的样本均值之和，即

$$\tau_1 + \tau_2 + \tau_3 = E(x_{11}) + E(x_{12}) + E(x_{13}) \tag{32}$$

并将 $\lambda_1 + \lambda_2 + \lambda_3$ 限制为第一阶段得到的负荷之和(基于合成总分标准化)或者将 $(\lambda_1 + \lambda_2 + \lambda_3)/3$ 限制为第一阶段得到的负荷均值(基于合成均分标准化)。显然，无论基于合成总分或者合成均分进行标准化，推出的限制等式都一样。

理论上说，这样标度的结果可以使得 η_1 成为标准化变量(均值为 0、方差为 1)，但实际上可能有出入，正如 Ferrer 等人(2008)说的这样得到的 η_1 均值近似于 0、方差近似于 1。因此，基于不同标度方法所做的标准化可能不完全一样。

3 一阶和二阶潜增长模型的可比性

对于多指标测量的变量，既可以使用合成分数建立一阶 LGM，也可以使用潜变量建立二阶 LGM。讨论一阶和二阶 LGM 的可比性，不仅在需要比较两者的时候能找对比较的对象，而且可以更好地理解二阶 LGM 标度方法。

3.1 基于尺度指标法的比较

由 2.1 节可知，对于使用尺度指标法的二阶 LGM，与仅使用尺度指标(而完全忽略其他指标)建立的单变量一阶 LGM 有可比性。具体来说就是，对 LGM 参数中的每一个，以 x_{t1} 为尺度指标的二阶 LGM 与用 x_{t1} 建立的单变量一阶 LGM 可以得到共同的方程，例如，$\mu_\alpha = E(x_{11})$，$\mu_\beta = E(x_{21}) - E(x_{11})$。但两者在 5 个 LGM 参数估计上还是会有出入，因为二阶 LGM 的这些参数估计会用到其他指标的信息。

这样说来，最常使用的尺度指标法并不理想，一方面，使用不同指标做尺度指标会得到可能很不同的结果，就如使用不同指标建立的单变量 LGM 会有很不同的结果一样。另一方面，使用尺度指标法，至少在初始水平和斜率的均值参数估计上主要依赖尺度指标，从而与效应编码法的结果可能会有

较大出入。

3.2　基于效应编码法的比较

由 2.2 节可知, 二阶 LGM 使用限制截距之和为 0、负荷之和为 1 的效应编码法, 与使用合成总分建立的单变量 LGM 有可比性, 两者可以得到主要 LGM 参数的 5 对一样的方程。但两种模型都属于超识别模型(侯杰泰 等, 2004), 关于参数的方程多于参数个数, 两种模型还有其他不一样的方程, 因而两种模型的参数估计只是接近而不是完全等同。和一般结构方程模型一样, 两种模型参数接近的程度与合成分数的信度有关(温忠麟 等, 2022)。

二阶 LGM 使用限制截距之和为 0、负荷之和为指标个数的效应编码法, 与使用平均分合成分数建立的单变量 LGM 有可比性, 两者可以得到主要 LGM 参数的 5 对一样的方程。两者关系的讨论同上。

3.3　基于潜变量标准化法的比较

理论上, 如果以 η_1 的标准化作为标度, 无论基于哪种标度方法去做 η_1 的标准化, 最后得到的标准化结果是一样的, 即二阶 LGM 的参数估计结果有唯一性。但实际上, 因为潜变量标准化法不是直接限制 η_1 的均值为 0、方差为 1, 基于不同标度方法去做 η_1 的标准化得到的结果可能不完全一致。而且, 截距因子均值的估计也可能不刚好等于 0。

与标准化的 η_1 建立的二阶 LGM 可比的单变量 LGM, 是用标准化合成分数 X_1 建模, 而且 X_2, X_3, X_4 要做与 X_1 相同的线性变换, 即减去 X_1 的均值后除以 X_1 的标准差, 使得 X_1 至 X_4 平移相同的距离后再做相同比例的尺度缩放。这是因为潜变量 η_1 至 η_4 在纵向测量强不变性限制下, 在 η_1 转换为标准化变量时, 都在相同距离的平移后有相同比例的缩放。因为合成总分和合成均分是倍数关系, 两者的标准化结果完全相同, 推出的单变量 LGM 也完全等同。

3.4　小结

根据上面的讨论, 可以将一阶和二阶 LGM 的比较结果列成表 1。对于多指标测量的变量, 要建立一阶 LGM 的时候, 几乎都会使用合成分数, 而罕有使用单指标建模; 但二阶 LGM 常用的潜变量标度方法却是尺度指标法(Yang et al., 2021)。例如, Leite (2007)通过模拟的方法比较一阶 LGM 和二阶 LGM 对参数的估计情况, 一阶 LGM 使用平均分合成分数建模, 而二阶 LGM 使用尺度指标法(并与产生数据的真模型一样), 两者缺乏可比性, 所以该文比较的结果没有意义。为了与使用平均分合成分数的一阶 LGM 比较, 二阶 LGM 应当使用平均分的效应编码法(并与产生数据的真模型一样)。

4　实际数据展示

这里用一个实例比较表 1 中各种标度方法的二阶 LGM 及相应的一阶 LGM, 看看 5 个 LGM 参数估计结果在不同标度方法上的差异, 以及一阶和二阶 LGM 可比性的情况。变量是道德推脱中的“责任转移”维度。这里主要是方法展示, 不过多涉及与方法无关的具体细节。

采用王兴超和杨继平(2010)修订的《中文版道德推脱问卷》。该问卷共 26 个题目, 采用 1 (完全不同意)~5(完全同意)五点计分, 共有 8 个维度, 选取其中的“责任转移”维度, 该维度有 3 个题目(指标)。被试是高校本科生, 追踪测量 4 次, 间隔 1 个月, 共有 1209 个被试有完整的数据。4 次测量的 α 系数在 0.72~0.87 之间。更多信息可参见金童林等(2023)。采用 Mplus 8.3 做 CFA(包括一阶和二阶 LGM)。

先做纵向测量不变性检验, 因为被试人数多, 不适合使用嵌套模型的卡方差异检验。改用拟合指数 CFI 和 RMSEA 的差异检验。对于人数超过 300、各组人数相当的多组测量不变性检验, 如果加上限制条件的模型的 CFI 降低不超过 0.01、RMSEA 提高不超过 0.015, 则选择比较简单的模型, 即不变性成立(Chen, 2007)。基准模型(没有限制条件)、弱不变性模型(限制各时间点的负荷相等)和强不变性模型(限制各时间点的负荷相等、截距相等)的主要拟

表 1　二阶 LGM 标度方法与可比的一阶 LGM 建模

二阶 LGM 标度方法	一阶 LGM	二阶 LGM	备注
尺度指标法	用尺度指标建模	尺度指标的截距固定为 0, 负荷固定为 1	一阶和二阶使用同一个尺度指标。
效应编码法	用合成分数建模	截距之和固定为 0, 负荷之和固定	一阶使用总分(或平均分), 二阶的负荷之和固定为 1(或指标个数)。
潜变量标准化法	参照点的合成分数标准化, 其他时间点的合成分数做相同的线性变换	基于尺度指标或者效应编码对参照点潜变量进行标准化	二阶 LGM 需要两阶段建模。对于同阶 LGM, 合成总分和合成均分得到的标准化结果相同。

表 2 纵向测量不变性检验

模型	χ^2	*df*	CFI	RMSEA
基准模型	652.4	48	0.925	0.102
弱不变性模型	670.8	54	0.923	0.097
强不变性模型	736.7	60	0.916	0.097

注：*df* = degrees of freedom, CFI = comparative fit index, RMSEA = root mean square error of approximation.

合指数见表 2。强不变性模型与基准模型相比，ΔCFI = −0.009 (CFI 降低不超过 0.01)，ΔRMSEA = −0.005 (RMSEA 不升反降)，所以纵向测量强不变性成立。

第一时间点潜变量均值为 0、方差为 1 时(Mplus 程序见附录 1)，3 个指标的负荷分别为：0.743, 0.679, 0.459，这些数值在潜变量标准化时需要用到(见 2.3 节)；相应的单指标信度分别为：0.67, 0.65, 0.25，可以通过信度高低检视尺度指标法的效果。

不同标度方法的二阶 LGM 及相应的一阶 LGM 的 5 个参数估计见表 3，为了方便比较，还加上了截距因子与斜率因子的相关系数，它不是 LGM 的参数，可以通过方差和协方差计算得到。不同标度方法的二阶 LGM 参数估计结果可能会很不相同，但模型的自由度都是 65 (可以用公式 23 计算)，而且拟合指数也相同：χ^2 = 754.52, CFI = 0.914, TLI = 0.913, RMSEA = 0.094, SRMR = 0.067。

对于尺度指标法，一阶和二阶 LGM 的参数估计结果有明显差异(通常最感兴趣的斜率因子均值差异超过 10%)，尤其是信度比较低的指标 3 作为尺度指标的时候。二阶 LGM 考虑了潜变量的测量误差，虽然主要依赖尺度指标，但还是使用全部指标的信息，而一阶 LGM 仅仅使用了尺度指标，而且没有考虑测量误差，当信度低(误差方差大)的时候，其估计结果与二阶 LGM 的差异大是意料中的事情。此外，无论一阶还是二阶 LGM，使用不同的尺度指标的结果可比性低。

对于效应编码法(二阶 LGM 的 Mplus 程序见附录 2)，无论合成总分还是合成均分建模，一阶和二阶 LGM 的参数估计差异都不超过 10%，两种模型都使用了全部指标信息，差异在于是否考虑了测量误差，即潜变量建模与显变量建模的差异。此外，无论一阶还是二阶 LGM，合成总分与合成均分建模尺度相差很大，但两者的增长参数符合理论预期的倍数关系。

当第一时间点的潜变量标准化时，可以得到 LGM 的标准化估计。理论上说，截距因子均值的标准化估计近似于 0。实际结果是，一阶 LGM 的截距因子均值近似于 0 且不显著；三种二阶 LGM 标准化估计，除了截距因子的均值外完全一致，但截距因子均值有明显差异，两种尺度指标法的结果与 0 有显著差异，而效应编码法的结果近似于 0 且不显

表 3 不同标度方法的二阶 LGM 及相应的一阶 LGM 比较结果

二阶 LGM 标度方法及其可比的一阶 LGM		截距因子均值(μ_α)	斜率因子均值(μ_β)	截距因子方差(σ_α^2)	斜率因子方差(σ_β^2)	截距因子与斜率因子的协方差($\sigma_{\alpha\beta}$)	截距因子与斜率因子的相关系数($r_{\alpha\beta}$)
尺度指标法	题目 1 (一阶)	2.11	−0.15	0.38	0.03	−0.04	−0.36
	题目 1 为尺度指标	2.06	−0.13	0.36	0.03	−0.03	−0.29
	题目 3 (一阶)	1.60	−0.04	0.37	0.03	−0.05	−0.50
	题目 3 为尺度指标	1.69	−0.08	0.14	0.01	−0.01	−0.29
效应编码法	合成总分(一阶)	5.66	−0.30	2.44	0.18	−0.19	−0.28
	负荷之和等于 1	5.72	−0.33	2.27	0.17	−0.18	−0.29
	合成均分(一阶)	1.89	−0.10	0.27	0.02	−0.02	−0.28
	负荷之和等于题数	1.91	−0.11	0.25	0.02	−0.02	−0.29
潜变量标准化法	合成总分(一阶标准化)	<u>−0.01</u>	−0.14	0.54	0.04	−0.04	−0.28
	合成均分(一阶标准化)	<u>−0.01</u>	−0.14	0.54	0.04	−0.04	−0.28
	基于题目 1 标准化	−0.11	−0.18	0.64	0.05	−0.05	−0.29
	基于题目 3 标准化	0.21	−0.18	0.64	0.05	−0.05	−0.29
	基于合成分数标准化	<u>0.02</u>	−0.18	0.64	0.05	−0.05	−0.29

注：没有注明一阶便是二阶。除了带下划线的三个截距因子均值(标准化估计)外，所有参数在 0.01 水平上都显著。不出所料，合成总分和合成均分的一阶 LGM 有完全相同的标准化估计。基于合成总分和合成均分的二阶 LGM 也有完全相同的标准化估计，没有分别列出，统称为基于合成分数的标准化估计。

著, 说明二阶 LGM 标准化应当使用效应编码法比较好。此外, 不出所料, 无论一阶还是二阶 LGM, 合成总分和合成均分建模有完全相同的标准化估计。

5 讨论

LGM 可以同时分析个体随时间的发展变化和个体间的发展差异(Jeon, & Kim, 2021; Laird & Ware, 1982), 在追踪研究中受到广泛应用。一阶 LGM 使用单变量建模, 容易理解、操作简单。我们以中国知网(https://www.cnki.net)全文数据库作为数据源, 不限出版年与学科, 检索所有涉及潜增长模型的中文文献。主题或摘要或关键词包括: "潜增长"、"潜增长模型"、"潜增长曲线模型"、"因子曲线模型"、"潜在增长"、"潜在增长模型"、"潜在增长曲线模型"、"LGM"。使用了 LGM 的应用文章 210 篇(其中心理学 119 篇), 仅有 5 篇使用了二阶 LGM (其中只有 1 篇心理学论文)。英文期刊上二阶 LGM 的使用也很少, Yang 等人(2021)检索了 4 个心理学期刊上从创刊至 2019 年 11 月的全部文章, 共有 300 篇文章应用了 LGM, 只有 11 篇使用了二阶 LGM。无论中文还是英文文章, 只有少数报告了使用尺度指标法, 多数没有报告标度方法。绝大多数文章使用 Mplus 进行二阶 LGM 分析, 可以推测是使用了默认的尺度指标法。

单变量一阶 LGM 的不足之处是, 忽略了测量误差、通常不去检验纵向测量不变性。对于多指标测量的变量, 使用合成分数做一阶 LGM 的优点是简单易理解, 前提条件也是满足纵向测量不变性并且是严格不变性(即各指标的负荷、截距和误差方差都是跨时间不变的)。如果严格不变性不成立, 使用合成分数做单变量 LGM 会引起参数估计偏差, 因为这种 LGM 不能区分潜变量的真实纵向变化与测量上的变化(Leite, 2007; Sayer & Cumsille, 2001)。二阶 LGM 可以弥补一阶 LGM 的不足, Geiser 等人(2013)总结了二阶 LGM 的优势: 可以将测量误差分离出来; 可以检验纵向测量不变性(其实这个检验可以独立于 LGM 进行); 在检验发展变化的个体差异方面有较高的统计检验力; 可以分离题目指标的方法效应。

正确使用二阶 LGM 的一个前提是理解模型识别和因子的标度方法, 以及二阶与一阶的联系。本文基于 Yang 等人(2021)的标度方法分类, 简化了尺度指标法和效应编码法的推演方式, 提出了简单而有可操作性的二阶 LGM 标准化估计方法, 系统总结了二阶 LGM 标度方法及其可比的一阶 LGM 建模。

除非对某个指标情有独钟而且该指标信度很高, 否则不要使用尺度指标法, 因为这样标度的二阶 LGM 结果只与尺度指标建立的单变量 LGM 有可比性, 不同的尺度指标可能导致很不同的结果。效应编码法比较好, 同时使用了全部指标的信息进行标度, 结果与使用合成分数的单变量 LGM 有可比性。可以根据研究目的决定使用总分还是平均分的效应编码法, 通常使用平均分的比较好, 不受指标数量的影响。

和其他结构方程模型一样, 标准化估计可以帮助解释结果, 比较不同研究之间的效应大小。虽然理论上说基于不同标度方法得到的标准化估计有唯一性, 但由于潜变量标准化的间接性, 标准化后的参照点潜变量只是近似的标准化变量, 因而参数估计可能与标准化方法有关。基于效应编码法的标准化明显好于基于尺度指标法的标准化。

由于不同标度方法的 LGM 结果很不相同, 需要明确说明所用的标度方法, 并对估计结果做出相应的解释。一般情况下推荐使用效应编码法对二阶 LGM 进行标度和标准化。这时, 有可比性的一阶 LGM 是使用合成分数的单变量 LGM。对一阶 LGM 而言, 合成总分和合成均分得到的标准化结果相同; 对二阶 LGM 而言, 合成总分和合成均分对应的效应编码法的标准化结果也相同。

表 3 列出的三种二阶 LGM 标准化估计, 除了截距因子均值外, 其他 4 个 LGM 参数估计都完全一致。这是巧合还是必然, 需要进一步研究。还有, 二阶 LGM 与一阶 LGM 的关系是潜变量建模与显变量建模的关系, 但比回归模型中的这种关系要复杂, 两者的关系如何受到测量信度的影响, 有待进一步研究。

参　考　文　献

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504.

Chou, C. P., Bentler, P. M., Pentz, M. A. (1998). Comparisons of two statistical approaches to study growth curves: The multilevel model and the latent curve analysis. *Structural Equation Modeling, 5*(3), 247–266.

Ferrer, E., Balluerka, N., & Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 4*(1), 22–36.

Geiser, C., Keller, B., & Lockhart G. (2013). First versus second order latent growth curve models: Some insights from latent state-trait theory. *Structural Equation Modeling,*

20(3), 479–503.

Grimm, K. J., Ram, N., & Estabrook, R. (2017). *Growth modeling: Structural equation and multilevel modeling approaches*. Guilford Press.

Hau K. T., Wen Z., & Cheng Z. (2004). *Structural Equation Model and Its Applications*. Beijing: Educational Science Publishing House.

[侯杰泰, 温忠麟, 成子娟. (2004). *结构方程模型及其应用*. 北京：教育科学出版社.]

Jeon, M. J., & Kim, S. Y. (2021). Performance of second-order latent growth model under partial longitudinal measurement invariance: A comparison of two scaling approaches. *Structural Equation Modeling*, *28*(2), 261–277.

Jin, T., Xu, M., & Wu, Y. T. N. (2023). The changes of moral disengagement in Chinese college students: Evidence from cross-temporal meta-analysis and longitudinal study. *Journal of Guizhou Normal University (Natural Sciences)*, *41*(1), 110–116.

[金童林, 徐明爽, 乌云特娜. (2023). 中国大学生道德推脱的变迁：横断历史研究与追踪研究的证据. *贵州师范大学学报(自然科学版)*, *41*(1), 110–116.]

Kim, E. S., & Willson, V. L. (2014). Testing measurement invariance across groups in longitudinal data: Multigroup second-order latent growth model. *Structural Equation Modeling*, *21*(4), 566–576.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38*(4), 963–974.

Leite W. L. (2007). A comparison of latent growth models for constructs measured by multiple items. *Structural Equation Modeling*, *14*(4), 581–610.

Li, L., Gao, Y., Zhang M., & Zhang, Y. (2012). The latent variable growth curve model and its applications. *Chinese Journal of Health Statistics*, *29*(5), 713–716.

[李丽霞, 郜艳晖, 张敏, 张岩波. (2012). 潜变量增长曲线模型及其应用. *中国卫生统计*, *29*(5), 713–716.]

Li, L., Zhou, S., Zhang, M., Zhang, Y., & Gao, Y. (2014). Comparisons of two statistical approaches in studying the longitudinal data: The multilevel model and the latent growth curve model. *Chinese Journal of Epidemiology*, *35*(6), 741–744.

[李丽霞, 周舒冬, 张敏, 张岩波, 郜艳晖. (2014). 多水平模型和潜变量增长曲线模型在纵向数据分析中的应用及比较. *中华流行病学杂志*, *35*(6), 741–744.]

Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling*, *13*(1), 59–72.

Mayer, A., Steyer, R., & Mueller, H. (2012). A general approach to defining latent growth components. *Structural Equation Modeling, 19*(4), 513–533.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561–614). New York: Plenum.

Sayer, A. G., & Cumsille, P. E. (2001). Second-order latent growth models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 179–200). Washington, DC: American Psychological Association.

Wang, X., & Yang, J. (2010). Reliability and validity of moral disengagement scale in Chinese students. *Chinese Journal of Clinical Psychology*, *18*(2), 177–179.

[王兴超, 杨继平. (2010). 中文版道德推脱问卷的信效度研究. *中国临床心理学杂志*, *18*(2), 177–179.]

Wen, Z., Chen, H., Fang, J., Ye, B., & Cai, B. (2022). Research on test reliability in China's mainland from 2001 to 2020. *Advance in Psychological Science*, *30*(8), 1682–1691.

[温忠麟, 陈虹熹, 方杰, 叶宝娟, 蔡保贞. (2022). 新世纪 20 年国内测验信度研究. *心理科学进展, 30*(8), 1682–1691.]

Wen, Z., & Liu, H. (2020). *Analyses of mediating and moderating effects: Methods and applications*. Beijing: Educational Science Publishing House.

[温忠麟, 刘红云. (2020). *中介效应和调节效应: 方法及应用*. 北京: 教育科学出版社.]

Wen, Z., & Ye, B. (2011). Evaluating test reliability: From coefficient alpha to internal consistency reliability. *Acta Psychologica Sinica, 43*(7), 821–829.

[温忠麟, 叶宝娟. (2011). 测验信度估计：从 α 系数到内部一致性信度. *心理学报*, *43*(7), 821–829.]

Wu, Y., & Wen, Z. (2011). Item parceling strategies in structural equation modelling. *Advance in Psychological Science*, *19*(12), 1859–1867.

[吴艳, 温忠麟. (2011). 结构方程建模中的题目打包策略. *心理科学进展*, *19*(12), 1859–1867.]

Yang, Y., Luo, Y., & Zhang, Q. (2021). A cautionary note on identification and scaling issues in second-order latent growth models. *Structural Equation Modeling*, *28*(2), 302–313.

附录 1：第一时间点潜变量标准化的因子分析 Mplus 程序(限制负荷和截距跨时间不变)

```
TITLE: CFA with strong MI (standardized f1)
DATA: FILE IS p1.csv;
VARIABLE: NAMES ARE x11-x13 x21-x23 x31-x33
          x41-x43;
MODEL:
      f1 BY x11-x13* (L1-L3);
      f2 BY x21-x23* (L1-L3);
      f3 BY x31-x33* (L1-L3);
      f4 BY x41-x43* (L1-L3);
    !设定负荷自由估计，跨时间不变;
        [x11 x21 x31 x41] (tau1);
        [x12 x22 x32 x42] (tau2);
        [x13 x23 x33 x43] (tau3);
    !设定截距跨时间不变;
      f1@1; !固定 f1 的方差为 1;
      [f1@0]; !固定 f1 的均值为 0, Mplus 默认;
      [f2-f4*]; !f2-f4 的均值自由估计;
!注释：结果用于二阶潜增长模型的潜变量标准化法(见附录 2);
```

附录 2：二阶潜增长模型的 Mplus 程序(效应编码法及其标准化)

```
TITLE: Second-order Latent Growth Models;
DATA: FILE IS p2.csv;
VARIABLE: NAMES ARE x11-x13 x21-x23 x31-x33
```

```
x41-x X1mean-X4mean Z1-Z4;
!Z1 为 X1mean 的 Z 分数, Z2-Z4 为 X2mean-X4mean 的线性变换;
USEVARIABLES are x11-x13 x21-x23 x31-x33 x41-x43;
MODEL:
f1 BY x11-x13* (L1-L3);
f2 BY x21-x23* (L1-L3);
f3 BY x31-x33* (L1-L3);
f4 BY x41-x43* (L1-L3);
!设定负荷自由估计, 跨时间不变;
[x11 x21 x31 x41] (tau1);
[x12 x22 x32 x42] (tau2);
[x13 x23 x33 x43] (tau3);
!设定截距跨时间不变;
i s | f1@0 f2@1 f3@2 f4@3;  !定义潜变量 LGM, i 为截距因子 α, s 为斜率因子 β;
[i*];  !截距因子均值自由估计;
i(ivar); s(svar); i with s (cov);
!为计算相关系数, 储存方差/协方差估计值;
MODEL CONSTRAINT:
NEW (rel);  !定义新变量;
rel=cov/sqrt(ivar*svar); !计算截距因子与斜率因子的相关系数;
0 = L1+L2+L3-3; !限制负荷之和等于指标个数(或者 1), 对应于合成均分(总分)建模;
0 = tau1+tau2+tau3; !限制截距之和等于 0;
!注释: 如果要得到标准化估计, 最后两行限制应当换成:
0 = L1+L2+L3-1.881;  !1.881 是第一阶段做 CFA (附录 1)得到的负荷之和;
0 = tau1+tau2+tau3-5.685;  !5.685 是指标在第一时间点的样本均值之和;
```

Scaling methods of second-order latent growth models and their comparable first-order latent growth models

WEN Zhonglin[1], WANG Yifan[1], DU Mingshi[1], YU Yahui[1], ZHANG Yuhui[1], JIN Tonglin[2]

([1] *Center for Studies of Psychological Application & School of Psychology*, *South China Normal University*, *Guangzhou* 510631, *China*)

([2] *College of Psychology*, *Inner Mongolia Normal University*, *Hohhot* 010022, *China*)

Abstract

Latent growth models (LGMs) are a powerful tool for analyzing longitudinal data, and have attracted the attention of scholars in psychology and other social science disciplines. For a latent variable measured by multiple indicators, we can establish both a univariate LGM (also called first-order LGM) based on composite scores and a latent variable LGM (also called second-order LGM) based on indicators. The two model types are special cases of the first-order and second-order factor models respectively. In either case, we need to scale the factors, that is, to specify their origin and unit. Under the condition of strong measurement invariance across time, the estimation of growth parameters in second-order LGMs depends on the scaling method of factors/latent variables. There are three scaling methods: the scaled-indicator method (also called the marker-variable identification method), the effect-coding method (also called the effect-coding identification method), and the latent-standardization method.

The existing latent-standardization method depends on the reliability of the scaled-indicator or the composite scores at the first time point. In this paper, we propose an operable latent-standardization method with two steps. In the first step, a CFA with strong measurement invariance is conducted by fixing the mean and variance of the latent variable at the first time point to 0 and 1 respectively. In the second step, estimated loadings in the first step are employed to establish the second-order LGM. If the standardization is based on the scaled-indicator method, the loading of the scaled-indicator is fixed to that obtained in the first step, and the intercept of the scaled-indicator is fixed to the sample mean of the scaled-indicator at the first time point. If the standardization is based on the effect-coding method, the sum of loadings is constrained to the sum of loadings obtained in the first step, and the sum of intercepts is constrained to the sum of the sample mean of all indicators

at the first time point. We also propose a first-order LGM standardization procedure based on the composite scores. First, we standardize the composite scores at the first time point, and make the same linear transformation of the composite scores at the other time points. Then we establish the first-order LGM, which is comparable with the second-order LGM scaled by the latent-standardization method.

The scaling methods of second-order LGMs and their comparable first-order LGMs are systematically summarized. The comparability is illustrated by modeling the empirical data of a Moral Evasion Questionnaire. For the scaled-indicator method, second-order LGMs and their comparable first-order LGMs are rather different in parameter estimates (especially when the reliability of the scale-indicator is low). For the effect-coding method, second-order LGMs and their comparable first-order LGMs are relatively close in parameter estimates. When the latent variable at the first time point is standardized, the mean of the intercept-factor of the first-order LGM is close to 0 and not statistically significant; so is the mean of the intercept-factor of the second-order LGM through the effect-coding method, but those through two scaled-indicator methods are statistically significant and different from each other.

According to our research results, the effect-coding method is recommended to scale and standardize the second-order LGMs, then comparable first-order LGMs are those based on the composite scores and their standardized models. For either the first-order or second-order LGM, the standardized results obtained by modeling composite total scores and composite mean scores are identical.

Keywords latent growth model, composite score, scaling method, scaled-indicator, effect-coding, latent-standardization